

# **AUTOMATED WORD PROCESSOR FOR CHINESE-STYLE LANGUAGES**

## **BACKGROUND OF THE INVENTION**

The background of this invention will be set forth in two parts.

### **1. FIELD OF INVENTION**

The present invention relates generally to an improved design of word and data processing device for the Chinese-style one-syllable words. This invention will facilitate an automated means for word and data processing involving the use of phonetically Latinized languages such as Chinese.

### **2. DESCRIPTION OF THE PRIOR ARTS**

Two types of word input devices for Chinese-style languages have been used. The first kind is to select individually each word from a memory reservoir of the Chinese-style vocabulary and to plant it at the intended location in the sequence of a sentence. The second kind is to type in phonetically each Chinese-style word in Latin or English alphabets for its sound value, then a corresponding word or words will appear on a display to be selected by the operator according to its meaning. Such as that:

2-01 yi (the sound value) = Chinese words: 一, 易, 乙, 医, 移, 亦, 伊, 宜, 姨, and more.

2-02 jiao (the sound value) = Chinese words: 叫, 脚, 教, 焦, 娇, 蕉, 交, 膠, and more.

There are also word-input devices that use the combination of these two methods.

The efficiencies of these devices are dismal. The first kind device is essentially the age old manual word-input machine. The second kind device cannot be used effectively because of the massive numbers of words in the Chinese-style languages that sound phonetically too close to be distinguishable from one another as shown in above Examples 2-01 and 2-02. The slow input speed is the basic cause of inefficiency of present Chinese-style word processors.

## **SUMMARY OF THE INVENTION**

The principal object of this invention is to provide means of high speed automated processing for words of Chinese-style languages. A new kind of data processing terminal devices can be constructed based on this invention. This input terminal adds a set of new sorting symbol keys to the traditional alphanumeric keys to produce classification effect of each Chinese word in Romanized pronunciation. Each Chinese word is produced phonetically with Roman alphabets; a Chinese sorting symbol is inserted before the Romanized Chinese word to define uniqueness of this individual word based on its specific meaning classification. This technique effectively eliminated the homonym problems of Chinese-type words. Chinese people are familiar with these selected distinctive symbols as they use similar ones every day in hand written Chinese characters.

It is also an object of this invention to produce a simple, effective, and reliable device that is also of low cost for high speed word and data processing.

Another purpose of this device is to further improve the efficiency in processing data with Chinese-style language input information. This can be accomplished by using data processing devices based on this invention in conjunction with communication software languages derived from the proprietary TRILAN "TM" vocabularies and its derivatives. The use of TRILAN "TM" will vastly facilitate the speed of communication among the three most widely used world languages, namely: English, Spanish, and Chinese.

## **BRIEF DESCRIPTION OF THE DRAWINGS**

Figure 1 is a schematic drawing indicating the functional relationship between this new improved terminal device and a data processor with output display. The input terminal device embodies the following elements:

1. Alphabetical letter keys of the Roman or English languages.
2. The numerical keys from 0 to 9.
3. Other miscellaneous symbolic and functional keys.
4. The specific Chinese character keys. These keys categorize meanings of Chinese words. Thus these keys provide a means for differentiating one Chinese word from many other Chinese words of similar phonetic values. A few of these are also used for stand-alone phonetic words as well. These distinctive symbols represent ancient first Chinese writing that were started from pictures. These symbols cannot be obtained by any simple combination of English alphabets. Specifically, these sorting symbols are:

- |      |  |
|------|--|
| 4-01 | 日 = a symbol indicating light, time                                |
| 4-02 | 文 = a symbol indicating culture, civilization                      |
| 4-03 | 方 = a symbol indicating direction, a state of being, a process     |
| 4-04 | 口 = a symbol indicating food, speak, actions associated with mouth |
| 4-05 | 人 = a symbol indicating people and human related activities        |
| 4-06 | 扌 = a symbol indicating action, movement                           |
| 4-07 | 力 = a symbol indicating the use of force                           |
| 4-08 | 电 = a symbol indicating phenomenon associated with electricity     |
| 4-09 | 耳 = a symbol indicating hearing, sound                             |
| 4-10 | 目 = a symbol indicating viewing with eyes                          |
| 4-11 | 身 = a symbol indicating body parts                                 |
| 4-12 | 手 = a symbol indicating actions associated with hand               |
| 4-13 | 艹 = a symbol indicating grass, vines, etc.                         |
| 4-14 | 木 = a symbol indicating wood, wood products                        |

- 4-15      氵 = a symbol indicating water, liquid
- 4-16      火 = a symbol indicating fire, heat
- 4-17      工 = a symbol indicating labor, industrial process
- 4-18      金 = a symbol indicating metal, metal products
- 4-19      土 = a symbol indicating soil, dirt, place
- 4-20      石 = a symbol indicating rock, mountain
- 4-21      纟 = a symbol indicating fabrics, organization
- 4-22      毛 = a symbol indicating wool, hair
- 4-23      𠂔 = a symbol indicating worship, respect
- 4-24      𠂔 = a symbol indicating covering
- 4-25      𠂔 = a symbol indicating mental activities
- 4-26      医 = a symbol indicating medical, health
- 4-27      之 = a symbol indicating transportation, destination, etc.
- 4-28      尸 = a symbol indicating building, rooms
- 4-29      虫 = a symbol indicating insects and smaller crawling animals
- 4-30      犴 = a symbol indicating four feet land animals
- 4-31      气 = a symbol indicating air, gas, aerial activities
- 4-32      羽 = a symbol indicating feather, flying activities
- 4-33      # = a symbol indicating measurements

This symbol 4-33, “#”, which appears in standard input keyboards is a Chinese character that had been used over 3,000 years. It means well (water well, gas well, etc.) in Chinese. It is pronounced as “jene” in Chinese. This symbol sounds similar with a fundamental unit in Chinese weight measurement. It is adopted here also as a Chinese sorting symbol

indicating general measurements classification. Additionally, the conventional symbol “\$” (meaning U. S. dollar) is adopted in this invention also as a sorting symbol for articles or actions associated with monetary activities when it is typed in as an integral part of a Latinized Chinese-style word.

## DETAILED DESCRIPTION OF THE INVENTION

Input Chinese words for common Chinese-style language processors now require operators to type in the sound value of these words spelled in Roman letters. Chinese government codified pronunciation of each of these words that rarely sound like ordinary conversational Chinese language. Since Chinese use only one-syllable words, the problem of homonyms becomes unbearable. None of these Romanized Chinese characters is unique, an operator has to select the correct Chinese character usually from a list of dozens of possibilities. This requirement of constant human intervention makes it unfit to achieve automation for the machines presently in use.

Chinese written words were invented based on both visual and sound effects together with a method of meaning categorization. Chinese written words are beyond the realm of pure phonetics. Chinese written words are least confusing as there are very few homonyms in them compare to all world languages that use alphabets. This was the reason historically to have all official commitments to be in writing. Using Roman letters a to z to provide the sound portion of Chinese written words, together with the special Chinese sorting symbols 4-01 through 4-33 and “\$” sign, Chinese words can be typed in with uniqueness. For example, referring to Articles 2-01 and 2-02, by writing: (The following samples are selected from copyrighted material TRILAN “TM”, a language compilation registered in United States of America, by Chang P. Liu. These

Romanized Chinese words follow closely with English and Spanish pronunciation customs. Chinese words selected in TRILAN “TM” sound close to daily conversational Chinese of a large region. These words are therefore easier to be remembered by users. These words are generally different from those officially codified words of China now in use for word and data processing.)

#yi = 一 (meaning the number “1” in Chinese and nothing else).

#yi-bei = 一百 (meaning the number “100” in Chinese and nothing else).

□jiao-hane = 叫喊 (meaning “yelling” in Chinese and nothing else).

\$jiao-yi = 交易 (meaning “trade” in Chinese and nothing else).

This one-to-one relationship between input Romanized words and corresponding output in genuine Chinese words indicates that the traditional requirement of human intervention can be eliminated or greatly reduced. This invention effectively increases input speed for Chinese-style words and data processing. It also increases processing reliability as the factor of human errors is reduced to minimum.

The Chinese sorting character symbols 4-01 to 4-32 will be designed using ASCII encoding system to be in compliance with international protocols. These symbols will function just as all other existing alphanumeric characters.

The claims of this invention are:

1. A data processing device wherein the improvement comprises:

(a) Element or elements that are capable of producing Chinese-style language characters based on a method for transmitting data involving Chinese-style language words, said method including steps of transmitting inputs including alphanumeric symbols of visual or audio characters, Chinese words sorting symbols and miscellaneous keys for system operation from a system input terminal and its elements to a data processor including memory elements for said inputs and other necessary functional elements for processing, and transmitting the product of the processor to an output terminal including compatible display for displaying; and

(b) A class of communication devices that is capable of processing the proprietary TRILAN "TM" vocabulary software and its derivative language compilations.

2. A class of data processing components that is capable of storing or processing in any way the information that involves the use of the Chinese sorting characters:

“日”, “文”, “方”, “口”, “亻”, “彳”, “力”, “电”, “尸”, “目”, “夕”, “扌”, “𠂇”, “木”, “彡”, “𠂆”, “工”, “金”, “土”, “石”, “纟”, “毛”, “衤”, “衤”, “巾”, “医”, “之”, “尸”, “虫”, “犭”, “气”, and “习” in combination with phonetically Latinized words of the Chinese-style words.

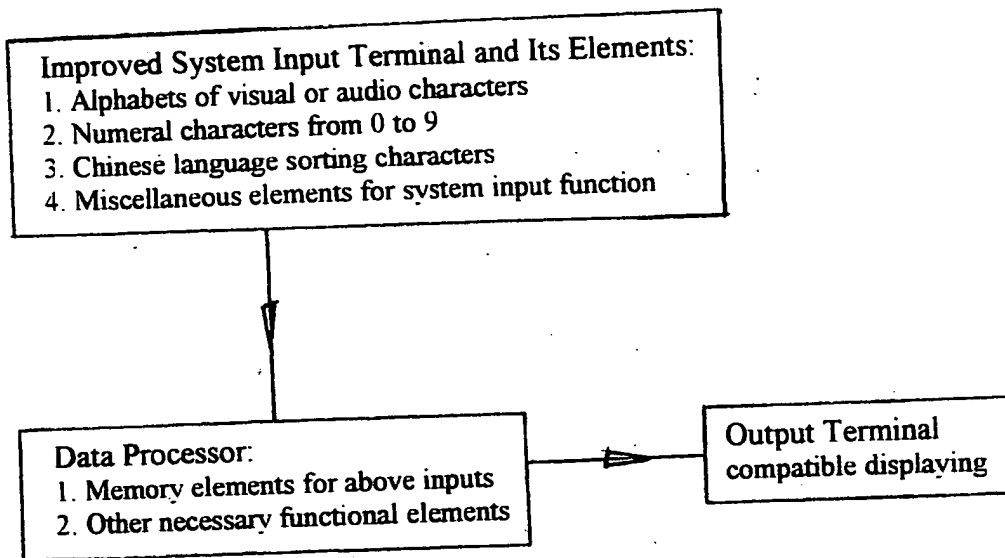


Figure 1. Schematic of the Improved Information Processing System for Transmitting Data That Involving Chinese-style Language Words.



日 文	方 口	イ イ	力 电
目 目	分 才	李 木	三 矢
工 金	土 石	毛 毛	木 木
十 医	之 戸	虫 牙	气 习

Figure 2. A Schematic Arrangement of the Chinese Character Keys on the Improvement Input Keyboard. (Actual arrangement may be different).